Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 3117–3121, May 1987
Biochemistry

Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria

(mitochondrial evolution/presequences/protein transport/DNA transposition)

ALISON BAKER AND GOTTFRIED SCHATZ

Department of Biochemistry, Biocenter, University of Basel, CH-4056 Basel, Switzerland

Communicated by Giuseppe Attardi, January 5, 1987 (received for review November 30, 1986)

ABSTRACT **Sequences that are capable of restoring mitochondrial targeting to a truncated yeast cytochrome *c* oxidase subunit IV presequence are encoded within the genome of *Escherichia coli* and within the gene for a higher eukaryotic cytosolic protein, mouse dihydrofolate reductase. These sequences, which resemble authentic presequences in their overall amino acid composition and degree of hydrophobicity, are rather frequent; >2.7% of clones generated from *E. coli* DNA and >5% of clones from the dihydrofolate reductase gene were functional in our screening system. These results suggest that, during evolution, mitochondrial precursor proteins could arise as a result of DNA rearrangements that place potential mitochondrial presequences at the amino terminus of existing open reading frames. Primitive eukaryotic cells may have used this mechanism to target proteins to their endosymbiotic protomitochondria.**

Many nuclear-encoded polypeptides destined for import into mitochondria are synthesized as precursors, bearing a transient amino-terminal presequence that is removed upon transport of the polypeptide to its correct location within the organelle (1). Experiments employing gene fusion have demonstrated that the information required for finding mitochondria and the correct submitochondrial compartment can be contained within the presequence (for review, see ref. 2). Presequences can thus act independently of the attached protein as signals for intracellular protein sorting.

How did mitochondrial presequences and the mechanism(s) for importing proteins into mitochondria evolve? It is generally accepted that mitochondria and chloroplasts are the descendents of formerly free-living prokaryotes that entered into symbiosis with other cell types (for review, see ref. 3). Mitochondria, however, depend on nuclear genes for over 90% of their proteins, including those required for the replication and expression of mitochondrial DNA. Assuming that the endosymbiont hypothesis for the evolution of eukaryotic cells is correct, many genes have moved from the endosymbiont to the nucleus, or have been lost altogether. This implies that mechanisms must have evolved to target the corresponding gene products to mitochondria, or to permit novel proteins to enter mitochondria. An intermediate stage may be represented by subunit 9 of F_0-ATPase, a nuclear gene that encodes a precursor with a presequence in *Neurospora crassa* and mammals; but in *Saccharomyces* the gene is located in the mitochondria, and the encoded polypeptide lacks a presequence (4). *Neurospora* mtDNA also contains a presequence-less copy of this gene, but whether it is expressed is unknown (5). Thus, the evolution of import pathways probably included the acquisition of amino-terminal presequences. This could have happened in at least two mutually compatible ways: (*i*) Presequences arose by the accumulation of point mutations in the 5′ untranslated sequences of genes, which generated additional coding sequences at the amino terminus. This may have occurred in the evolution of the yeast mitochondrial isozyme of alcohol dehydrogenase (6). (*ii*) Presequences were created by DNA rearrangements that added additional blocks of amino acids to the amino termini of existing proteins.

In order to address the latter question, we asked whether sequences present within prokaryotic DNA, or a eukaryotic cytosolic protein, could restore mitochondrial targeting to a nonfunctional cytochrome *c* oxidase subunit IV precursor lacking most of its presequence.

MATERIALS AND METHODS

Strains and Culture Conditions. *Saccharomyces cerevisiae* strains WD1 (*alpha, his3, ura3, coxIV*) (7) and CF1 (*alpha, his3, ura3, coxIV*) (C. Fankhauser, unpublished work) were used. CF1 and WD1 strains are derived from the same strain, DL1 (8), but strain CF1 differs from strain WD1 in that it has a more extensive deletion of the subunit IV gene that includes the entire presequence—extending from the upstream *Eco*RI site to the *Cla* I site within the region encoding the mature subunit (9). The culture media were minimal medium (MM; 0.67% yeast nitrogen base, 2% glucose) supplemented with histidine and, where necessary, with uracil and leucine each at 10 μg/ml, YPD (1% yeast extract/2% peptone/2% glucose), and YPEG (1% yeast extract/2% peptone/3% ethanol/3% glycerol). *Escherichia coli* strain GM99 (*F-, dam4, mal354, tsX354*) was obtained from T. Bickle (Biocenter, University of Basel, Switzerland), and was grown in Luria broth (1% tryptone/0.5% yeast extract/1% NaCl).

Construction of Presequence Clones. DNA was extracted from *E. coli* (10), purified by equilibrium centrifugation in a CsCl gradient, and partially digested with DNase I (11), or completely digested with a combination of the restriction enzymes *Alu* I, *Hae* III, and *Rsa* I. Fragments of ≈20–70 base pairs (bp) migrating between the marker dyes xylene cyanol and bromophenol blue were purified from an 11% polyacrylamide gel (12). The ends of these fragments were made blunt with T4 DNA polymerase and ligated into plasmid pYCX11 which had been linearized by digestion with *Xba* I, rendered blunt-ended by the action of the "Klenow fragment" of *E. coli* DNA polymerase I and then treated with calf intestinal phosphatase. The 660-bp *Bam*HI/*Hin*dIII fragment containing the coding region for mouse dihydrofolate reductase (DHFR) and short stretches of flanking sequence was purified from plasmid pDS5/2 (13). The purified fragment was incubated with DNase I, such that fragments of <70 bp were generated. The ends of the DNase I-digested fragments were made blunt by consecutive treatment with T4 DNA polymer-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: DHFR, dihydrofolate reductase.

ase and the Klenow fragment of *E. coli* DNA polymerase I and cloned into the filled-in and phosphatased *Xba* I site of pYCX11. Recombinant plasmids were used to transform *E. coli* strain RR1 to ampicillin resistance. Plasmid DNA was extracted from six pools of recombinants—each pool representing the products of individual ligation reactions and the plasmid DNA was then used to transform subunit IV-deficient yeast cells. In the case of the DHFR clones, DNA was extracted from two pools containing ≈4000 colonies per pool. Yeast transformants were first selected on uracil-free MM plates for acquisition of the plasmid-borne *ura3* marker and then replica-plated to YPEG plates to score for respiring transformants (7).

Miscellaneous Methods. Recombinant DNA methods were as described (14). Published methods were used for the transformation of yeast (15), extraction of proteins from whole yeast cells, and for immune blotting (16). Plasmids were isolated from yeast transformants by a modification of the alkaline lysis method used for *E. coli* (14), in which cells harvested from logarithmic phase cultures were treated with Zymolyase 20,000 (1.5 mg/ml) (Kirin Brewery, Japan) in 1 M sorbitol/0.1 M sodium citrate/0.06 M Na_2EDTA, pH 7/0.5% beta-mercaptoethanol for 60 min at 37°C, in place of the usual lysozyme/EDTA treatment. DNA sequence analysis was performed by the dideoxy-chain termination method (17) using M13 vectors (18).

Analysis of Presequence Clones. A set of theoretical "random" sequences was generated from the DHFR gene as follows. Fifteen random sites within the DHFR gene were taken as the starting point of hypothetical subclones. Fragments encoding 15 amino acids (including the fusion points) were "ligated" into the same vector, pYCX11, as was used in the real experiment. When stop codons in the sequence prevented the insertion of more than nine consecutive amino acids, the initially selected starting point was shifted by plus or minus one. For each of the 15 start points 5'-to-3' sequences were generated from both DNA strands, yielding a total of 30 sequences. These sequences were analyzed in the same way as the import-competent sequences selected in the screening experiments. Hydrophobic moments and average hydrophobicities were calculated according to ref. 19.

RESULTS

The Screening System. Sequences that transport cytochrome oxidase subunit IV into yeast mitochondria can be detected from among a collection of essentially random sequences by a functional assay based on the complementation of a subunit IV-deficient mutant by a plasmid-encoded copy of this gene (20). The DNA sequences are inserted into the truncated subunit IV presequence of the plasmid pYCX11 (plasmid d3-20, ref. 21). This yeast–*E. coli* shuttle plasmid contains a unique *Xba* I site between the 3rd and 20th codons for the truncated presequence. Fusion proteins were expressed and tested for their ability to restore the growth of the subunit IV-deficient mutant on ethanol and glycerol. We refer to these selected sequences as "import-competent," and these sequences together with the seven amino acids from the subunit IV presequence as "composite functional presequences."

***E. coli* DNA Contains Many Sequences That Can Restore Import Competence to a Truncated Presequence.** *E. coli* DNA was digested with DNase I, or a combination of restriction enzymes, and fragments of 20–70 bp were isolated and tested for their ability to reconstitute a functional mitochondrial presequence. Of the 9160 yeast transformants obtained, 246 (2.7%) could grow on glycerol and had, thus, acquired a functional presequence for the truncated subunit IV. This percentage (2.7%) underestimates the true frequency of complementing sequences for the following reasons. (*i*) Only one-third of the fragments should have inserted such that the start codon is in frame with the cytochrome oxidase subunit IV gene. (*ii*) Some of the correctly inserted fragments may also contain in-frame stop codons. (*iii*) Some of our 9160 transformants may harbor plasmids containing a very short, or even no, insert of *E. coli* DNA.

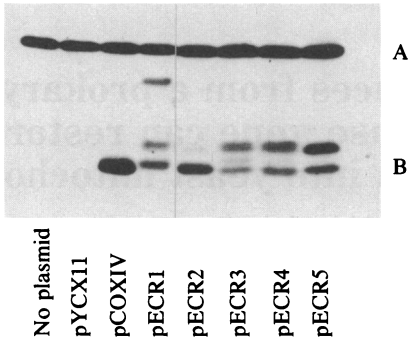

FIG. 1. Subunit IV-related proteins in respiring yeast transformants containing an *E. coli*-derived subunit IV presequence. The subunit IV-deficient mutant CF1, transformed with the indicated plasmids, was grown in MM media lacking uracil to an OD_{600} of ≈1. Total cellular proteins were extracted from 2 ml of the cultures, separated by electrophoresis in a 13% $NaDodSO_4$/polyacrylamide gel, and subjected to immune blotting with antisera against cytochrome *c* oxidase subunit IV (B) and the major 29-kDa pore-forming protein (porin) of the yeast mitochondrial outer membrane (A). This latter protein was tested to allow comparisons between the amount of protein in each sample. Bound immunoglobulin G was visualized with iodinated protein A. An autoradiograph of the nitrocellulose replica is shown.

Five of the respiring yeast transformants were studied in more detail. When these plasmids were recovered and retransformed into the subunit IV-deficient mutant, the mutant cells again became respiration competent (data not shown). In order to identify the subunit IV polypeptides made by the respiring transformants, the five transformants were grown on glucose-containing media, and the subunit IV-related polypeptides were detected by immune blotting (Fig. 1).

As expected, no subunit IV could be detected in the mutant cells themselves (identified in the figure as "no plasmid"). This was also true in this autoradiograph for mutant cells transformed with the vector pYCX11, because the truncated subunit IV precursor made by these cells cannot enter mitochondria and is unstable in the cytosol (20, 21); however, in general, some subunit IV antigen is visible in longer autoradiographic exposures. Each of the transformants carrying an inserted piece of *E. coli* DNA accumulated a larger-sized subunit IV and, in addition, at least one subunit IV-related polypeptide of higher mobility. The smaller bands may result from cleavage by the mitochondrial-processing enzyme or by nonspecific proteolysis of the presequence (21–23). This experiment does not indicate whether the assembled cytochrome oxidase of these transformants contains only mature-sized cytochrome oxidase subunit IV or also the larger forms. All samples shown in Fig. 1 contain similar amounts of mitochondrial porin, a protein of the mitochondrial outer membrane, indicating that the total amount of mitochondrial protein is similar in each case.

The sequences of the *E. coli* DNA fragments that restore mitochondrial targeting are shown in Fig. 2. We compared these sequences to *E. coli* sequences from the GenBank DNA sequence library* but were unable to determine their origin.

Both Coding and Noncoding Regions of DNA Can Specify Import-Competent Sequences. In order to determine more

*National Institutes of Health (1985) Genetic Sequence Databank: GenBank (Research Systems Div., Bolt, Beranek, and Newman, Cambridge, MA), Tape Release 38.0.

ECR1 M L S T I T T R Q T T R V A V T V V V A L P V F R A V E H H V F Q H
L T Q T R T A F R W H Q F F F N R R Q R E Q S R Y L L

ECR2 M L S R L P C C A A R K D A L A S L T L A R Y L L

ECR3 M L S S L S T P W R A A W F S L R A R Y L L

ECR4 M L S S R S L I N A A F T M I T R A R Y L L

ECR5 M L S R L I I T A S A L R N S V S R G P S W L S S F I D V S T F F Q
F W D V S G S L P V L P P A R Y L L

FIG. 2. Amino acid sequences of the presequences derived from *E. coli* DNA. The fusion genes encoding subunit IV with an *E. coli*-derived presequence were transferred to the M13 vectors mp9 and mp10 as *Eco*RI/*Hin*dIII or *Eco*RV/*Eco*RI fragments and sequenced (17, 18). Amino acids derived from the authentic subunit IV presequence are underlined. Basic residues are circled; acidic residues are boxed.

precisely where potential mitochondrial targeting sequences might be found, we turned to the gene encoding a cytosolic protein from a higher eukaryote, mouse DHFR. A 660-bp cDNA fragment containing the entire protein-coding sequence and a small amount of flanking sequence was digested with DNase I, and the fragments were treated as described in the preceding sections. We recovered 11,000 yeast transformants; 550 of these (5%) grew on glycerol/ethanol and had thus acquired a functional presequence for the plasmid-encoded subunit IV. The sequences of 27 of these clones were determined. Twenty of these sequences were unique, and their positions within the DHFR sequence were located. As shown in Fig. 3, they came from all six possible reading frames within this sequence.

Import-Competent Sequences Derived from *E. coli* or the Mouse DHFR-Encoding Gene Are Similar to Natural Mitochondrial Presequences. The import-competent sequences derived from DHFR (Fig. 4) were compared to a sample of 30 random sequences of similar length generated theoretically from the 660-bp restriction fragment that contains the DHFR-coding sequence. The random sequences, as expected, have essentially the same amino acid composition as the parental DHFR molecule (Table 1, columns B and C). However, the import-competent sequences (column A) show a marked bias against the negatively charged residues aspartic and glutamic acid. This paucity of negatively charged amino acids is a general feature of natural mitochondrial presequences.

Also characteristic of natural mitochondrial presequences is a high proportion of positively charged residues. Table 1 does not reveal selection for fragments with a higher than average proportion of positive charges. However, if the contributions of the Arg residue present in the carboxyl-terminal portion of the truncated subunit IV presequence and the Arg created at the 5′ fusion point in 50% of the sequences are considered, then the percentage of Lys plus Arg increases to 13.5—lower than the average value but within the range of natural presequences.

As hydroxylated residues are often abundant in natural presequences, we also calculated the mol % of Ser and Thr in our import-competent and random sequences; no significant differences were found. However, DHFR itself has a much higher-than-average content of Ser, and this may mask any selection for this residue in our import-competent sequences.

These competent sequences also cluster within a narrower range of average hydrophobicity (−0.23 to +0.34) than the random group (−0.52 to +0.57) including, in each case, the contribution of the seven amino acids of the subunit IV presequence. The average hydrophobicities of the ten natural mitochondrial presequences from Table 1 range from (−0.25 to +0.06). Also striking is that 80% of the functional composite sequences, compared with 90% of the natural presequences but only 40% of the random composite sequences, contain a region with a hydrophobic moment of 0.5 or greater. These data show that the import-competent sequences recovered from DHFR form a specific subset of the total sequences.

DISCUSSION

Because the *E. coli* genome is so large (≈4500 kb) it is reasonable to assume that our complementation system selected import-restoring sequences from an essentially random pool of DNA sequences. While this is not the case for the presequences derived from the DHFR gene, both selections yielded composite presequences that were similar to one another and to natural presequences. Although we cannot exclude cooperativity between the remaining parts of the yeast cytochrome oxidase subunit IV presequence and our inserted *E. coli* or mouse DNA fragments, we feel that our "import-competent" sequences are playing an active role in the restoration of a functional presequence, rather than acting as inert spacers that merely present the residual parts of the subunit IV presequence to the import machinery. The deficiency of acidic amino acids in our inserted fragments, the average amino acid composition, and the overall neutral hydrophobicity of the functional composite presequences argue that a specific subset of all available sequences was selected. The contribution of the seven amino acids from the subunit IV presequence remains unclear. The mitochondrial import machinery does not recognize a specific amino acid sequence, as functional synthetic presequences can be built using only three different kinds of amino acids (21). This implies that the role of the residual amino acids from the subunit IV presequence must be limited to participating in, or

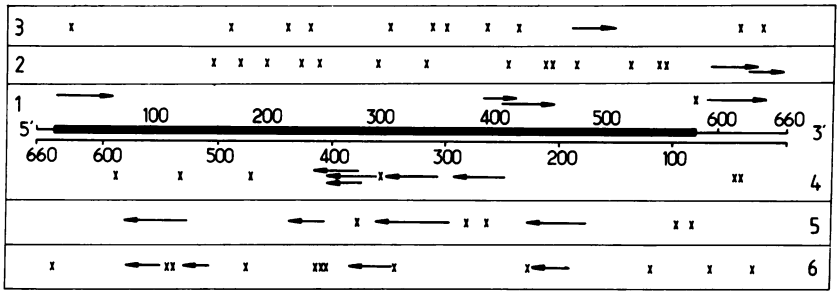

FIG. 3. Location of functional mitochondrial presequences within the DHFR-encoding gene. The six possible translational reading frames are numbered 1 to 6. Reading frame 1 encodes the DHFR polypeptide: the open reading frame on the 660-bp *Bam*HI/*Hin*dIII fragment is indicated on the map as a solid black box. ×, stop codons in the various reading frames; black arrows, origin of the DNA inserts recovered from the genes encoding import-competent artificial subunit IV precursors.

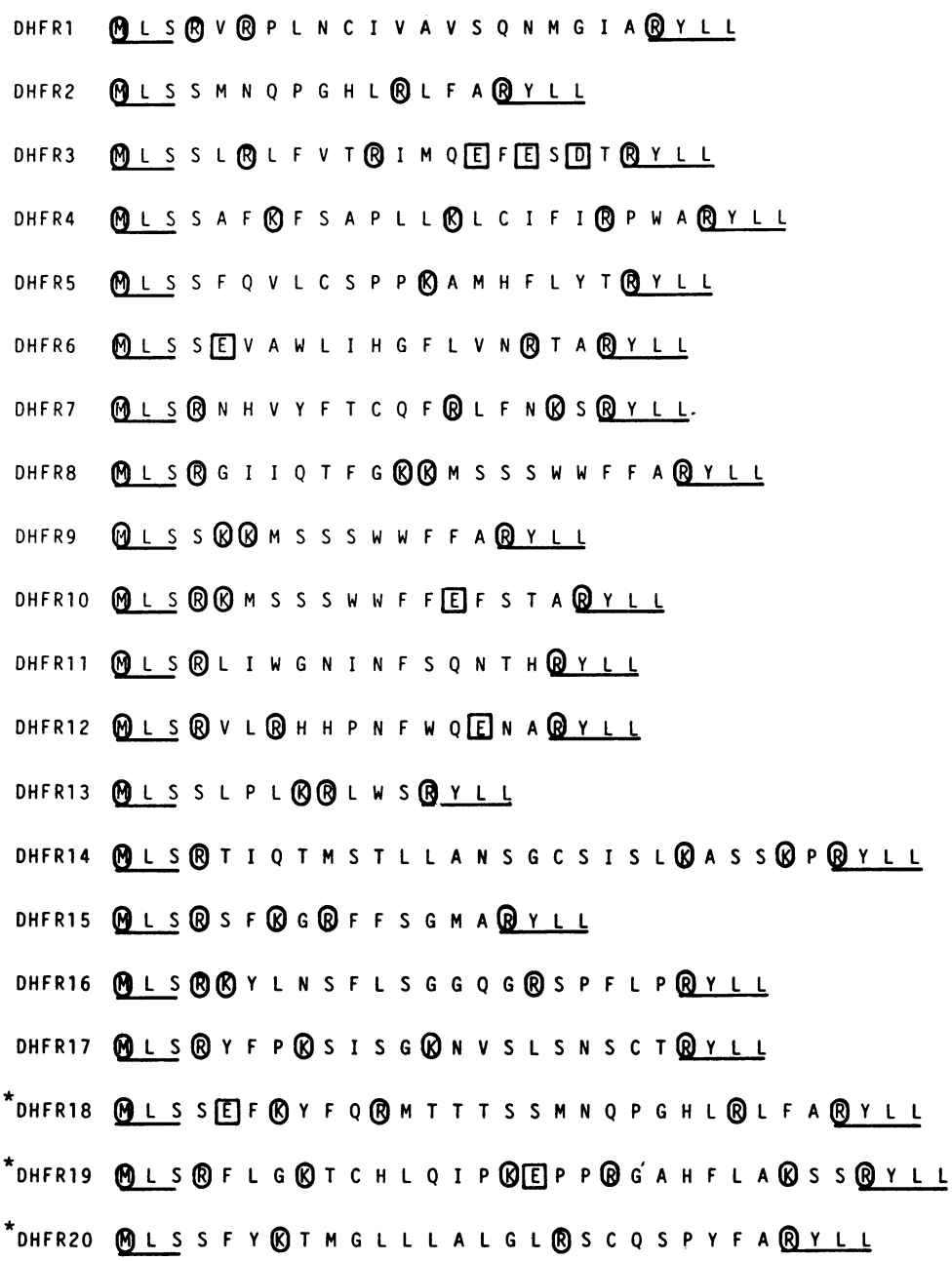

FIG. 4. Amino acid sequences of the functional presequences recovered from mouse DHFR. Amino acids derived from the cytochrome oxidase subunit IV presequence are underlined. Basic residues are circled, and acidic residues are boxed. *, Compound clones generated by the ligation of two noncontiguous DNA fragments from DHFR-encoding gene.

stabilizing, some essential higher-order structure. Furthermore, the carboxyl-terminal half of the natural presequence (including the sequence Arg-Tyr-Leu-Leu present in our constructs) can be deleted with little effect on import (33).

Work with chemically synthesized mitochondrial prepeptides and theoretical predictions (34, 35) have suggested that mitochondrial presequences may form amphiphilic structures upon interaction with membranes. In support of this, an α-helical region with amphiphilic character that is normally located in the interior of mouse DHFR was shown to function as a mitochondrial presequence when placed in front of "passenger" proteins (36). Interestingly, this helical region was not detected in our screen of randomly generated fragments. This may be because we have not analyzed sufficient clones to find all possible functional sequences, although the recovery of duplicate clones and overlapping fragments suggests that we are approaching saturation of the DHFR gene. An alternative explanation is that this sequence may contain a site that is unusually sensitive to DNase I. The experiments described here located a potential mitochondrial targeting sequence in the extreme amino-terminal region of DHFR, which had been scored as inactive in the other study. However, differences in the constructs generate an extra positive charge at the amino terminus in our clone and remove a negative charge and a stretch of hydrophilic residues. Even single amino acid changes, such as the removal of an acidic residue or the deletion of two proline residues, were found to have a significant effect on the presequence function (36).

Most of the sequences generated in this study also exhibited some amphiphilic character when analyzed (19, 34). However, the relatively small size of our sample and the assumptions required in calculating the amphiphilicity of such peptides lead us to interpret these results with caution. An additional complication in such analysis is that only part of each presequence may be essential for import. In the case of some natural presequences, the functional region has been localized to the first 12 (33) or 9 (37) residues, whereas in another case it appeared to be close to the center of the presequence (38). While such flexibility on the part of the

Table 1. Import-competent sequences are deficient in negatively charged amino acids

	A	B	C	D	E
Arg	4.9%*	7.8%	7.5%	4.5%	12.2%
Lys	6.5%	6.7%	5.9%	6.6%	5.7%
Glu	2.2%	5.5%	5.1%	5.6%	0.0%
Asp	0.3%	2.5%	3.0%	5.5%	0.0%

The amino acid composition (mol %) of import-competent sequences (column A) was compared with those of the random sequences derived from the DHFR gene (column B), with all reading frames in the DHFR gene (column C), with the average for a sample of 185 proteins (column D) (24), and with the average for 10 mitochondrial presequences (column E). The mitochondrial presequences were the authentic presequences of yeast citrate synthase (25), yeast cytochrome oxidase subunit IV (9), yeast cytochrome oxidase subunit V (26), yeast cytochrome oxidase subunit VI (27), yeast manganese superoxide dismutase (28), yeast alcohol dehydrogenase mitochondrial isozyme (6), yeast β subunit of F_1-ATPase (29), the first 32 residues of yeast cytochrome c_1 (30), and the rat (31) and human (32) ornithine carbamoyltransferase presequences.

*Inclusion of the arginine residues from the truncated yeast cytochrome oxidase subunit IV presequence and from the fusion points increases this value to 9.8.

import machinery may have facilitated the evolution of import pathways, it necessitates caution in interpreting structural parameters of presequences—particularly if these parameters have been derived by averaging over the entire length of the presequence. Nevertheless, the functional presequences isolated in this study share several general features with natural presequences.

Central to this study is the demonstration that prokaryotic genomes contain a large repertoire of sequences capable of restoring mitochondrial targeting to a "disabled" presequence. Most likely, sequences with independent targeting function also exist. Our analysis of the DHFR gene shows that these sequences do not necessarily originate from protein-coding regions, but may also come from alternate reading frames or noncoding regions. Such potential targeting sequences are normally inactive, either because they are not expressed as polypeptides, or because they are masked within the folded proteins. It is attractive to speculate that such sequences played an important role in the evolution of protein-sorting pathways within primitive eukaryotic cells. These sequences could have become activated by DNA transposition, a process we have mimicked here by the use of recombinant DNA technology.

Statistical analysis of protein coding genes allows identification of independently evolved open reading frames (40). When this method was applied to the gene encoding the cytochrome oxidase subunit IV precursor, it strongly suggested that presequence and "mature" part had been derived from separate genes. While such a dual origin was not detected for other mitochondrial precursors, the observation concerning the subunit IV precursor is consistent with the results presented here. Although no evolutionary pathway can be conclusively proven in the laboratory, our results provide one plausible explanation for the origin of presequences.

Note Added in Proof. Kaiser *et al.* (39) have recently shown that a high percentage of random fragments of human DNA can provide a secretory signal for yeast invertase, and Vassarotti *et al.* (41) have reported that point mutations in the amino terminus of the mature F_1-ATPase subunit can generate a functional mitochondrial import sequence.

The skillful and willing help of Kitaru Suda, who did the DNA sequencing, and of Wolfgang Oppliger, who made the photographs, is gratefully acknowledged. John Shepherd set up the sequence-homology searches for us, Christoph Fankhauser constructed the CF1 deletion strain, and Michèle Probst typed the manuscript. Many useful suggestions came from colleagues in the "Schatz Lab." A.B. is supported by a Long-Term Fellowship from the European Molecular Biology Organization. This study was supported by Grants 3.394-0.83 and 3.660-0.84 from the Swiss National Science Foundation.

1. Hay, R., Boehni, P. & Gasser, S. (1984) *Biochim. Biophys. Acta* **779**, 65–87.
2. Hurt, E. C. & van Loon, A. P. G. M. (1986) *Trends Biochem. Sci.* **11**, 204–207.
3. Gray, M. W. & Doolittle, W. F. (1982) *Microbiol. Rev.* **46**, 1–42.
4. Viebrock, A., Perz, A. & Sebald, W. (1982) *EMBO J.* **1**, 565–571.
5. van den Boogaart, P., Samallo, J. & Agsteribbe, E. (1982) *Nature (London)* **298**, 187–189.
6. Young, E. T. & Pilgrim, D. (1985) *Mol. Cell. Biol.* **5**, 3024–3034.
7. Dowhan, W., Bibus, C. R. & Schatz, G. (1985) *EMBO J.* **4**, 179–184.
8. van Loon, A. P. G. M., van Eijk, E. & Grivell, L. A. (1983) *EMBO J.* **2**, 1765–1770.
9. Maarse, A. C., van Loon, A. P. G. M., Riezman, H., Gregor, I., Schatz, G. & Grivell, L. A. (1984) *EMBO J.* **3**, 2831–2837.
10. Marmur, J. (1961) *J. Mol. Biol.* **3**, 208–218.
11. Anderson, S. (1981) *Nucleic Acids Res.* **9**, 3015–3027.
12. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4835–4838.
13. Stueber, D., Ibrahimi, I., Cutler, D., Dobberstein, B. & Bujard, H. (1984) *EMBO J.* **3**, 3143–3148.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
15. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. (1983) *J. Bacteriol.* **153**, 163–168.
16. Yaffe, M. P. & Schatz, G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4819–4823.
17. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3643–3646.
18. Messing, J. (1983) *Methods Enzymol.* **101**, 20–78.
19. Eisenberg, D., Schwarz, M., Komaromy, M. & Wall, R. (1984) *J. Mol. Biol.* **179**, 125–142.
20. Hurt, E. C., Müller, U. & Schatz, G. (1985) *EMBO J.* **4**, 3509–3518.
21. Allison, D. S. & Schatz, G. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9011–9015.
22. Boehni, P., Gasser, S., Leaver, C. & Schatz, G. (1980) in *Organisation and Expression of the Mitochondrial Genome*, eds. Kroon, A. M. & Saccone, C. (Elsevier/North-Holland Biomedical, Amsterdam), pp. 423–433.
23. McAda, P. C. & Douglas, M. G. (1982) *J. Biol. Chem.* **257**, 3177–3182.
24. Dayhoff, M. O., Dayhoff, R. E. & Hunt, L. T. (1976) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 2.
25. Suissa, M., Suda, K. & Schatz, G. (1984) *EMBO J.* **3**, 1773–1781.
26. Koerner, T. J., Hill, J. & Tzagoloff, A. (1985) *J. Biol. Chem.* **260**, 9513–9515.
27. Wright, R. M., Ko, C., Cumsky, M. G. & Poyton, R. O. (1984) *J. Biol. Chem.* **259**, 15401–15407.
28. Marres, C. A. M., van Loon, A. P. G. M., Oudshoorn, P., van Steeg, H., Grivell, L. A. & Slater, E. C. (1985) *Eur. J. Biochem.* **147**, 153–161.
29. Takeda, M., Vassarotti, A. & Douglas, M. G. (1986) *J. Biol. Chem.* **261**, 10466.
30. Sadler, I., Suda, K., Schatz, G., Kaudewitz, F. & Haid, A. (1984) *EMBO J.* **3**, 2137–2143.
31. McIntyre, P., Graf, L., Mercer, J. F. B., Wake, S. A., Hudson, P. & Hoogenraad, N. (1985) *DNA* **4**, 147–156.
32. Horwich, A. L., Wayne, A. F., Williams, K. R., Kalousek, F., Kraus, J. P., Doolittle, R. F., Konigsberg, W. & Rosenberg, L. E. (1984) *Science* **224**, 1068–1074.
33. Hurt, E. C., Pesold-Hurt, B., Suda, K., Oppliger, W. & Schatz, G. (1985) *EMBO J.* **4**, 2061–2068.
34. von Heijne, G. (1986) *EMBO J.* **5**, 1335–1342.
35. Roise, D., Horvath, S. J., Tomich, J. M., Richards, J. H. & Schatz, G. (1986) *EMBO J.* **5**, 1327–1334.
36. Hurt, E. C. & Schatz, G. (1987) *Nature (London)* **325**, 499–503.
37. Keng, T., Alani, E. & Guarente, L. (1986) *Mol. Cell. Biol.* **6**, 355–364.
38. Horwich, A. L., Kalousek, F., Fenton, W. A., Pollock, R. A. & Rosenberg, L. E. (1986) *Cell* **44**, 451–459.
39. Kaiser, C. A., Preuss, D., Grisafi, P. & Botstein, D. (1986) *Science* **235**, 312–317.
40. Shepherd, J. C. W. (1986) *Chem. Scr.* **26B**, 75–83.
41. Vassarotti, A., Stroud, R. & Douglas, M. (1987) *EMBO J.* **6**, 705–711.